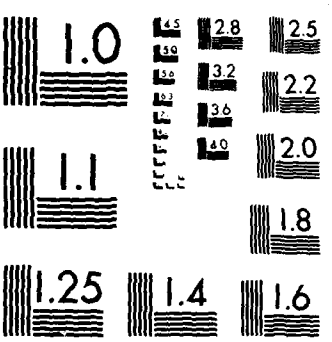
AD-A175 829 DISTRIBUTION OF THE MAXIMUM OF A GAUSSIAN PROCESS BY MONTE CARLO(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROC A M HASOFER JUL 86 TR-147 1/1

UNCLASSIFIED AFOSR-TR-86-2147 F49620-85-C-0144 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

DISTRIBUTION OF THE MAXIMUM OF A GAUSSIAN PROCESS BY MONTE CARLO

by

A.M. Hasofer

Technical Report No. 147

July 1986

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY NA	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 147	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 86-2147
6a. NAME OF PERFORMING ORGANIZATION University of North Carolina 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM
6c. ADDRESS (City, State and ZIP Code) Center for Stochastic Processes, Statistics Department, Phillips Hall 039-A, Chapel Hill, NC 27514	7b. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) nm	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620 85 C 0144
8c. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, DC	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. 6.1102F; PROJECT NO. 2304; TASK NO. A5; WORK UNIT NO.
11. TITLE (Include Security Classification) Distribution of the maximum of a Gaussian process by Monte Carlo	
12. PERSONAL AUTHOR(S) Hasofer, A.M.	
13a. TYPE OF REPORT technical	13b. TIME COVERED FROM 9/1/85 TO 9/30/86
14. DATE OF REPORT (Yr., Mo., Day) July 1986	15. PAGE COUNT 29
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES FIELD / GROUP / SUB. GR. XXXXX XXXXXXXXXXXX	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) ~~Keywords:~~

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

First a simple practical procedure for approximating a stationary Gaussian process over a finite interval by a trigonometric polynomial with predetermined error is described.

The approximation is then used to calculate the distribution of the maximum, using a novel Monte Carlo method with a control variable which drastically reduces the variance.

Finally, the outlined approach is compared to the moving-average technique and shown to be superior for continuous-time, narrow-band processes.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ OTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL ~~Peggy Rawitch~~ Capt Thomas	22b. TELEPHONE NUMBER (Include Area Code) ~~919-962-2307~~ 767-5025 22c. OFFICE SYMBOL AFOSR/NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

Distribution of the maximum of a Gaussian Process by Monte Carlo

by

A.M. Hasofer

Department of Statistics
University of New South Wales
Kensington, NSW, Australia

and

Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

ABSTRACT

First a simple practical procedure for approximating a stationary Gaussian process over a finite interval by a trigonometric polynomial with predetermined error is described.

The approximation is then used to calculate the distribution of the maximum, using a novel Monte Carlo method with a control variable which drastically reduces the variance.

Finally, the outlined approach is compared to the moving-average technique and shown to be superior for continuous-time, narrow-band processes.

1. Introduction

Many problems associated with stochastic processes in continuous time cannot be solved exactly by analytical methods. The essential reason for this lies, in the author's opinion, in the fact that these stochastic processes, even over a finite time interval, consist of a non-denumerable infinity of dependent variables. It is on account of this that ever since the advent of the modern computer, researchers using stochastic processes have resorted to simulation to obtain the results they required.

However, in spite of the widespread use of simulation, there has been an extraordinary dearth of theoretical work in this area.

While the fundamental principles of simulation of continuous-time stochastic processes were formulated by Rice (1954), the standard exposition of the simulation approach based on the spectrum has been given by Shinozuka and Jan (1972) and Yang (1972,1973).

Since then, there have been swift developments in the area of simulation of Markov-type and queueing stochastic processes, which are used extensively in Operations Research. See for instance the special issue of "Operations Research", Vol. 31, No. 6 (Nov. 1983), dedicated to simulation.

On the other hand there have been only a few scattered papers on simulation of continuous-time stochastic-processes based on the spectrum, the most easily accessible of which are Bílý and Bukoveczky (1976), Mihailov (1978) and Malyshev and Palagin (1981). See also Bílý and Čačko (1982) and Kropáč (1981).

All proposed simulation schemes rely on approximating the stochastic process by a function of a finite number n of random variables. This is almost invariably justified by showing that the approximation converges in some sense to the original stochastic process as n tends to infinity.

However, this type of justification gives no indication of the size of the error committed when n is taken to be a comparatively small number. In many applications the size of this error is not very crucial, because the original process is not known with a high precision anyway.

But in the application addressed to in this paper, namely the numerical calculation of a distribution function, given an exact formula for the spectrum, it is of fundamental importance to know the size of the error.

In this paper, two new ideas are put forward to obtain a computational technique which is more efficient than those previously proposed, and for which exact bounds for the error can be evaluated.

The first idea can be applied to any stationary Gaussian process $X(t)$ with given spectral density $f(\lambda)$ and consists in representing the process over a finite interval by a finite trigonometric polynomial in such a way that the error committed can be exactly bounded.

The second idea relates to the calculation of the distribution function of the maximum. The finite trigonometric approximation is used to simulate the process. The maximum functional is then split up into two independent components. The distribution of one component is calculated exactly, while that of the other is obtained from the simulation. The two components are then combined, yielding a high precision estimate for the distribution of the maximum.

The germ of the last idea is to be found in Deak (1980), but its application to the calculation of the distribution of the maximum of a stochastic process was given independently by the author in Hasofer (1982). See also Moran (1984).

The importance of using well-defined bounds for the approximation error is further highlighted by the results given in Lyon (1970), where it is shown that approximations which might be satisfactory for low levels might be completely useless for studying the behaviour of the stochastic process at high levels.

2. Construction of approximating polynomial

In Hasofer (1982), it is shown that, over a finite interval $(-T, +T)$, a real, zero-mean, Gaussian stationary stochastic process $X(t)$ with spectral density $f(\lambda)$ can be approximated by a finite random trigonometric sum $W_N(t)$, in the sense that there exist two stochastic processes $W_N'(t)$ and $Y(t)$, such that $X(t) + Y(t)$ and $W_N(t) + W_N'(t)$ are

identical in probability law on $(-T, +T)$. Using the symbol $\stackrel{d}{=}$ for equality in probability law, we have $X(t) + Y(t) \stackrel{d}{=} W_N(t) + W_N'(t)$. Moreover $X(t)$, $Y(t)$, $W_N(t)$ and $W_N'(t)$ may be chosen so as to be independent.

Through a judicious choice of T and the terms of $W_N(t)$, the two processes $W_N'(t)$ and $Y(t)$ can be made to be very "small", thus allowing $W_N(t)$ to approximate $X(t)$.

The procedure for constructing $W_N(t)$, $W_N'(t)$ and $Y(t)$ is as follows:

(a) Let $$R(t) = \int_{-\infty}^{+\infty} e^{it\lambda} f(\lambda) d\lambda \quad \ldots\ldots\ldots\ldots \quad (1)$$

be the covariance function of $X(t)$. Expand $R(t)$ in a Fourier series over $(-2T, +2T)$, in the form

$$R(t) = \sum_{n=0}^{\infty} \alpha_n \cos \frac{n\pi t}{2T} \ .$$

Let $a_n^2 = \max(0, \alpha_n)$, $b_n^2 = \max(0, -\alpha_n)$.

(b) Choose a minimal subset of integers, N, such that

$$\sum_{n \in \bar{N}} a_n^2 \ll \sum_{n \in N} a_n^2,$$

where $\bar{N}$ is the complement of N in the set of non-negative integers.

(c) Then

$$W_N(t) = \sum_{n\in N} a_n \left(V_n \cos \frac{n\pi t}{2T} + W_n \sin \frac{n\pi t}{2T}\right) ,$$

$$W_N'(t) = \sum_{n\in N} a_n \left(V_n \cos \frac{n\pi t}{2T} + W_n \sin \frac{n\pi t}{2T}\right) ,$$

$$Y(t) = \sum_{n} b_n \left(V_n \cos \frac{n\pi t}{2T} + W_n \sin \frac{n\pi t}{2T}\right) ,$$

where the $\{V_n\}$ and the $\{W_n\}$ are sequences of independent Gaussian random variables with zero mean and unit variance.

The notion of "smallness" in this context refers to the two coefficients σ_o and σ_2 defined for the process $Z(t)$ by

$$\sigma_o^2 = \text{Var}\,[Z(t)]$$

and

$$\sigma_2^2 = \text{Var}\,[Z'(t)] ,$$

since $\Pr\{\max_{-T\le t\le T} |Z(t)| > x\} \le 2\{1-\Phi(x/\sigma_o)\} + (2T\sigma_2/\pi\sigma_o)\exp[-\frac{1}{2}(x^2/\sigma_o^2]$

...(2)

where $\Phi(x)$ is the standard normal cumulative distribution function and the right-hand side of equation (2) is an increasing function of both σ_o and σ_2, for $x > \sigma_2$ (see Hasofer 1982).

It is to be noted that by taking N sufficiently large, $W_N'(t)$ can be made as small as required, while this is not the case with $Y(t)$.

The process $Y(t)$ is derived from the negative coefficients of the Fourier expansion of $R(t)$ over the finite interval $(-2T, +2T)$. It is well known that, as $T \to \infty$, the Fourier series tends to the corresponding Fourier integral with non-negative integrand. Thus, $Y(t)$ can be made as small as required, by taking T sufficiently large.

As an example of the application of the above idea, one can approximate the distribution of the maximum of $X(t)$ over $(-T, +T)$ by that of $W_N(t)$ as follows:

Let $$M_X = \max X(t)$$

$$M_1 = \max W_N(t)$$

$$U(t) = Y(t) + W_N'(t)$$

$$M = \max |U(t)|$$

where the maximum is taken over all values of t in $(-T, +T)$.

It is then easy to establish the following inequality, which holds for every k.

$$P(M_1>x+2k) - 2P(M>k) \le P(M_X>x) \le P(M_1>x-2k) + 2P(M>k)$$

(see Hasofer (1982)).

Clearly the error depends on $P(M > k)$, which in turn is bounded as in (2). Thus a measure of the error is the magnitude of the two coefficients

$$\text{Var}\,[U(t)] = \sum_n b_n^2 + \sum_{n\in N} a_n^2 ,$$

and

$$\text{Var}\,[U'(t)] = \frac{\pi^2}{4T^2}\left(\sum_n n^2 b_n^2 + \sum_{n\in N} n^2 a_n^2\right).$$

3. Relation to discrete approximation of distributions

Looking at the spectrum of the process $X(t)$, which has been assumed to be continuous (equation (1)), and that of the process $W_N(t)$, which is discrete, when $W_N(t)$ is continued to $(-\infty, +\infty)$ by periodicity, we see that the approximation developed will produce (apart from a scale factor) a discrete approximation to the absolutely continuous probability density function $f(\lambda)/R(0)$.

Such approximations have been shown to be very useful in many practical problems, for example in the evaluation of moments of functions of random variables. (See e.g. the much quoted paper of Rosenlueth (1975)). The procedure described in principle in Section 2 and detailed in the sequel provides a simple algorithm for obtaining a discrete approximation with many optimal properties which can be derived from the optimality properties of Fourier polynomials.

4. Optimal property of the proposed representation

Let $W^*(t)$ be another approximation to $X(t)$, with covariance function $R^*(t)$. Expanding $R^*(t)$ in a Fourier Series over the interval $(-2T, +2T)$, we have, say

$$R^*(t) = \sum_{n=0}^{\infty} \beta_n \cos\frac{n\pi t}{2T}$$

For simplicity let us write e_n for $\cos(n\pi t/2T)$, x^+ for $\max(0,x)$ and x^- for $-\min(0,x)$.

Then
$$R(t) = \sum_{n=0}^{\infty} \alpha_n^+ e_n - \sum_{n=0}^{\infty} \alpha_n^- e_n.$$

and
$$R^*(t) = \sum_{n=0}^{\infty} \beta_n^+ e_n - \sum_{n=0}^{\infty} \beta_n^- e_n.$$

It follows that

$$R(t) + \sum \alpha_n^- e_n = R^*(t) + \sum (\alpha_n^+ - \beta_n) e_n$$
$$= R^*(t) + \sum (\alpha_n^+ - \beta_n)^+ e_n - \sum (\alpha_n^+ - \beta_n)^- e_n$$

or

$$R(t) + \sum \alpha_n^- e_n + \sum (\alpha_n^+ - \beta_n)^- e_n = R^*(t) + \sum (\alpha_n^+ - \beta_n)^+ e_n.$$

This leads to the representation

$$X(t) + Y(t) + Y^*(t) \stackrel{d}{=} W^*(t) + Z^*(t),$$

where $Y^*(t)$ has covariance functicn $\sum (\alpha_n^+ - \beta_n)^- e_n$ and $Z^*(t)$ has covariance function $\sum (\alpha_n^+ - \beta_n)^+ e_n$.

Measuring the discrepancy between $X(t)$ and $W^*(t)$ as in the previous section, we see that both coefficients

$$\mathrm{Var}[\underline{Y}(t) + \underline{Y}^*(t) + Z^*(t)]$$

$$\mathrm{Var}[\underline{Y}'(t) + \underline{Y}^{*\prime}(t) + Z^{*\prime}(t)]$$

are minimized by choosing $\beta_n = \alpha_n^+$, which implies that $W^*(t)$ has the same distribution as $W_N(t) + W_N'(t)$. Finally a standard least squares argument shows that the trigonometric polynomial of order N which gives the smallest error as defined above is indeed $W_N(t)$.

5. Some practical formulae for evaluation of the coefficients and the remainder terms

For the purpose of determining the appropriate set N, as well as the value of T, it is best to start with the spectral density function $f(\lambda)$ of $X(t)$. This is an even function of λ and

$$R(t) = \int_{-\infty}^{+\infty} \cos(\lambda t) f(\lambda)\, d\lambda .$$

The Fourier coefficients of $R(t)$ in $(-2T, +2T)$ are given by

$$\alpha_o = \frac{1}{2T} \int_o^{2T} R(t)dt \quad ,$$

$$\alpha_n = \frac{1}{T} \int_o^{2T} R(t) \cos \frac{n\pi t}{2T}\, dt \qquad (n > o),$$

and these can be rewritten as

$$\alpha_o = \int_{-\infty}^{+\infty} \frac{\sin 2\lambda T}{2\lambda T} f(\lambda)d\lambda \quad ,$$

$$\alpha_n = \int_{-\infty}^{+\infty} \left[\frac{\sin(2\lambda T - n\pi)}{(2\lambda T - n\pi)} + \frac{\sin(2\lambda T + n\pi)}{(2\lambda T + n\pi)} \right] f(\lambda) d\lambda, \quad (n > o).$$

This last formula can be rewritten more conveniently as

$$\alpha_n = (-1)^n \int_{-\infty}^{+\infty} \frac{4\lambda T \sin 2\lambda T}{(4\lambda^2 T^2 - n^2 \pi^2)} f(\lambda) d\lambda , \quad (n > o) . \tag{3}$$

One of the most interesting cases is that of a "narrow band" stochastic processes, where $f(\lambda)$ is of the form

$$f(\lambda) = \left[\frac{1}{2\sigma} \; h\left(\frac{\lambda_o - \lambda}{\sigma}\right) + h\left(\frac{\lambda_o + \lambda}{\sigma}\right) \right] ,$$

and $h(\lambda)$ is a density function with mass concentrated in a neighbourhood of the origin.

Formula (3) can then be rewritten as

$$\alpha_n = (-1)^n \int_{-\infty}^{+\infty} \frac{4\lambda T \sin 2\lambda T}{(4\lambda^2 T^2 - n^2 \pi^2)} h(z) dz, \quad n > o$$

with $\lambda = \sigma z + \lambda_o$,

and $\alpha_o = \int_{-\infty}^{+\infty} \frac{\sin 2\lambda T}{2\lambda T} h(z) dz,$

with λ as above.

To choose appropriate values for N and T, we notice that the distance between the angular velocities of two successive terms such as α_n and α_{n+1} is $\pi/2T$. Suppose that most of the mass of the spectral density function on the positive axis is concentrated in the interval $(\lambda_o-k\sigma,\lambda_o+k\sigma)$, where k is some real positive number. We now attempt to approximate $f(\lambda)$ by means of $2M+1$ masses equidistantly placed on the interval $(\lambda_o-k\sigma,\lambda_o+k\sigma)$. We must then have

$$\frac{2k\sigma}{2M} = \frac{\pi}{2T}$$

or

$$2T = \frac{\pi M}{k\sigma}$$

The middle index N_o will then be given by

$$N_o = \left[\frac{2T\lambda_o}{\pi} \right] = \left[\frac{M\lambda_o}{k\sigma} \right] ,$$

where $[\ ]$ denotes as usual the integral part, and the set of indices chosen for N will be

$$N_o-M,\ N_o-M+1,\ \ldots,\ N_o+M.$$

It will then be necessary to check that the values of α_n obtained are not negative. Any negative coefficients will have to be removed. The values of σ_o and σ_2 for any of the three processes $W_N(t)$, $W_N'(t)$, $Y(t)$ will be given by

$$\sigma_o^2 = \sum|\alpha_n| ,$$

$$\sigma_2^2 = \frac{\pi^2}{4T^2} \sum n^2|\alpha_n| ,$$

where the summation extends over the appropriate set of indices.

In practice, provided $f(\lambda)$ decays sufficiently fast as $\lambda \to \infty$, only a comparatively small number of coefficients will be necessary to adequately represent the covariance function, and then σ_o^2 and σ_2^2 can be evaluated for each of the three processes by using only the small set of non-negligible coefficients.

6. A numerical example

As an example of the application of the above methods, we consider a narrow-band process with the spectral density function

$$f(\lambda) = \frac{1}{2\sigma} \left[\phi\left(\frac{\lambda_o - \lambda}{\sigma}\right) + \phi\left(\frac{\lambda_o + \lambda}{\sigma}\right) \right]$$

where ϕ is the density function of the standard normal distribution

We choose the following parameters

$$\lambda_o = 15\pi \ ,$$

$$\sigma = 2\pi \ .$$

We then choose $2T = 1$ and $k = 5$, which gives us $M = 2Tk\sigma/\pi = 10$. The index corresponding to λ_o will then be $N_o = 2T\lambda_o/\pi = 15$.

Applying the above formulae, with numerical integration performed by Simpson's rule over a z-interval of (-7, +7) with 500 intervals, we obtain the coefficients for $W_N(t)$ given in Table 1.

Coefficients beyond n = 35 were negligible, and there were no significant negative coefficients, so that Y(t) in this case was negligible. The values of σ_o^2 and σ_2^2 for X(t) were $\sigma_o^2 = 1$ and

$$\sigma_2^2 = \lambda_o^2 + \sigma^2 = 222\pi^2 \cong 2260,$$

while the corresponding values for $W_N'(t)$ were

$$\sigma_o^2 = 8.58 \times 10^{-7}$$

and $$\sigma_2^2 = 4.98 \times 10^{-3},$$

showing that the approximation obtained with twenty coefficients is high.

TABLE 1

Fourier Coefficients for $W_N(t)$

n	$\alpha_n = a_n^2$	n	$\alpha_n = a_n^2$
5	7.43×10^{-7}	15	0.199
6	7.99×10^{-6}	16	0.176
7	6.69×10^{-5}	17	0.121
8	4.36×10^{-4}	18	6.48×10^{-2}
9	2.22×10^{-3}	19	2.69×10^{-2}
10	8.76×10^{-3}	20	8.76×10^{-3}
11	2.70×10^{-2}	21	2.22×10^{-3}
12	6.48×10^{-2}	22	4.36×10^{-4}
13	0.121	23	6.69×10^{-5}
14	0.176	24	7.99×10^{-6}

7. Calculation of $P(M_1 > x)$

It was pointed out in Hasofer (1982) that the distribution of the maximum of $W_N(t)$ can be calculated directly by means of a multiple integral. If there are k elements in N, the multiple integral extends over the $2k$-dimensional ball of unit radius.

To perform the integration, however, one cannot use efficiently any deterministic method. As was pointed out by Davis and Rabinowitz (1975), the Monte Carlo method is the preferred approach for integration in spaces of large dimensions.

In the light of the preceding paragraph, one can recast the problem in a purely probabilistic framework, as follows:

We have

$$W_N(t) = \sum_{n\epsilon N} a_n\left(V_n \cos\frac{n\pi t}{2T} + W_n \sin\frac{n\pi t}{2T}\right).$$

Write $R^2 = \sum_{n\epsilon N}(V_n^2 + W_n^2)$. Then R^2 has clearly a chi-squared distribution of order $2k$. We now note that R is independent of the random variables

$$A_n = \frac{V_n}{R}, \qquad B_n = \frac{W_n}{R}.$$

We thus have

$$W_N(t) = R\, T_N(t).$$

Where $$T_N(t) = \sum_{n \in N} a_n \left(A_n \cos\frac{n\pi t}{2T} + B_n \sin\frac{n\pi t}{2T}\right).$$

It follows that

$$M_1 = \max_{-T \le t \le +T} W_N(t) = RM_1^* ,$$

where $$M_1^* = \max_{-T \le t \le +T} T_N(t).$$

It finally follows that

$$P(M_1 > x) = P(RM_1^* > x)$$

$$= E\left[P\left(R > \frac{x}{M_1^*} \,\middle|\, M_1^*\right)\right] = E\left[P\left(\frac{R^2}{2} > \frac{x^2}{2M_1^*} \,\middle|\, M_1^*\right)\right]$$

which is the same as the formula given at the bottom of p. 341 in Hasofer (1982), since $R^2/2$ has a Gamma(k) distribution.

We now propose the following procedure for practical calculation of $P(M_1 > x)$:

(i) Generate $2\Omega k$ independent standard normal random variables $V_n^{(i)}$, $W_n^{(i)}$ $(i = 1, \ldots, \Omega)$; and calculate

$$R^{(i)} = \left(\sum_{n \in N} V_n^{(i)2} + W_n^{(i)2}\right)^{1/2}$$

(ii) Find the maximum $M_1^{*(i)}$ of the trigonometric polynomial

$$T_N^{(i)}(t) = \frac{1}{R^{(i)}} \sum_{n\epsilon N} a_n (V_n^{(i)} \cos\frac{n\pi t}{2T} + W_n^{(i)} \sin\frac{n\pi t}{2T})$$

(iii) Estimate $P(M_1 > x)$ by

$$\hat{P}(M_1 > x) = \frac{1}{\Omega} \sum_{i=1}^{\Omega} P(R > \frac{x}{M_1^{*(i)}} \mid M_1^{*(i)})$$

8. Properties of the estimator $\hat{P}$

(1) Clearly $\hat{P}$ is an unbiased estimator of $P(M_1 > x)$ since

$$E(\hat{P}) = \frac{1}{\Omega} \sum_{i=1}^{\Omega} E\left[P(R > \frac{x}{M_1^{*(i)}} \mid M_1^{*(i)})\right] = E\left[P(R > \frac{x}{M_1^*} \mid M_1^*)\right]$$

(2) For large Ω, $\hat{P}$, tends to be normally distributed, on account of the central limit theorem.

(3) The variance of $\hat{P}$ can easily be estimated. Let

$$P_i = P(R > \frac{x}{M_1^{*(i)}} \mid M_1^{*(i)})$$

Then an unbiased estimator of the variance of P_i is

$$S^2(P_i) = \frac{1}{(\Omega-1)} \sum_{i=1}^{\Omega} (P_i - \hat{P})^2,$$

and the corresponding unbiased estimator of the variance of $\hat{P}$ is

$$S^2(\hat{P}) = \frac{S(P_i)}{\Omega}$$

By taking a sufficiently large number of simulations, the variance of the estimator can be made as small as required.

The technique of separating out R^2 and using its exact distribution amounts to a "control-variable" variance reduction technique in Monte-Carlo integration. See e.g. Rubinstein (1981).

9. Some numerical results

The technique described in the preceding section was applied to the trigonometric polynomial given in Section 5, and the following results were obtained.

(1) M_1^*:

Mean: 0.327

Standard deviation: 0.077

(2) Tail of M_1

x	$\hat{P}(M_1 > x)$	$\sqrt{S^2(P_i)}$
4	1.82×10^{-3}	1.8×10^{-2}
5	9.68×10^{-6}	2.03×10^{-4}
6	2.917×10^{-9}	6.22×10^{-8}

(3) Number of simulations required.

Suppose we require a coefficient of variation of 5% for $\hat{P}$. Then the number of simulations required turns out to be:

x	Ω
4	40,000
5	175,900
6	182,700

These numbers are well within the capabilities of an average computer.

(4) Computer time per simulation

The simulation was carried out both on a large main frame computer (Control Data Cyber) and on a small personal computer (IBM PC XT). The average time per simulation for the main frame was 1.2 seconds, while the average time for the IBM PC was 18 seconds.

However, there was no difficulty in obtaining a large number of simulations on the PC, since it could be left running over week-ends, and for each week-end about 10,000 simulations could be accumulated.

10. Advantage of the above technique

One might think that with the above large number of simulations, it might not be necessary to go through all the steps of Section 8. One might be tempted to simulate $W_N(t)$ directly and estimate the tail probabilities of M_1 directly from its histogram. To appreciate the large gain effected by the procedure of Section 8, it is only necessary to calculate the number of simulations necessary to achieve the same coefficient of variation as in Section 9, namely 5%, by direct simulation.

The variance of the estimator $\hat{p}$ of a probability from a histogram is given by

$$s^2(\hat{p}) = \frac{p(1-p)}{\Omega^*}$$

where Ω^* is the number of simulations. From this formula we obtain

x	$\Omega^*(\approx \frac{400}{p})$
4	220,000
5	4.13×10^7
6	1.4×10^{11}

11. Approximation to the distribution of the maximum of X(t)

As mentioned above the following inequality holds

$$P(M_1>x+2k)-2P(M>k)\leq P(M_X>x)\leq P(M_1>x-2k)+2P(M>k).$$

Thus we can approximate the distribution of the tail of the maximum of $X(t)$ over $(-T, +T)$ by the distribution of the tail of the maximum of $W_N(t)$, calculated in Section 6.

The term $P(M > k)$ can be bounded by using formula (2).

Suppose we take $k = 0.005$. It then turns out that

$$P(M > k) < 10^{-5}.$$

Choose $x = 3$. It then turns out that

$$P(M_1 > 3.01) = 0.0624$$
$$P(M_1 > 2.99) = 0.0653$$

Thus

$$0.0624 \leq P(M_X > 3) \leq 0.0654$$

corresponding to a margin of error of $\pm$ 3%.

12. The modulus of variation of $T_N(t)$

One of the great advantages of the trigonometric polynomial representation is that upper bounds for the variability of the approximation can be easily obtained. The approach will be illustrated by the calculation of M_1^*, based on the maximum of a finite set of values of t.

Let

$$M_1^{**} = \max T_N(t_i)$$

where the t_i form a mesh of span Δ in $(-T, +T)$.

Let t_0 be the value of t for which $T_N(t)$ reaches M_1^*, and let $t_i \le t_o \le t_{i+1}$.

Now at t_0 we must have $dT_N|dt = 0$ with probability one. However

$$T_N(t_i) = T_N(t_0) + (t_i - t_0)T_N'(t_0) + \frac{(t_2 - t_0)^2}{2!} T_N''(t'),$$

where $t_1 \le t' \le t_0$.

Thus

$$T_N(t_0) - T_N(t_i) = \frac{(t_i - t_0)^2}{2!} [-T_N''(t')],$$

and since obviously $T_N(t_i) \le M_1^{**} \le T_N(t_0)$,

we have

$$|M_1^* - M_1^{**}| \le \frac{\Delta^2}{2} M_0$$

where

$$M_0 = \max_t |-T_N''(t)|.$$

But $-T_N''(t) = (\frac{\pi}{2T})^2 \sum_{n\in N} n^2 a_n (A_n \cos\frac{n\pi t}{2T} + B_n \sin\frac{n\pi t}{2T})$,

so that

$$M_0 = \max_t |-T_N''(t)| \le (\frac{\pi}{2T})^2 \sum_{n\in N} n^2 a_n \sqrt{A_n^2 + B_n^2} ,$$

$$\le (\frac{\pi}{2T})^2 (\sum_{n\in N} n^4 a_n^2)^{1/2}$$

using Schwarz's inequality and the fact that $\sum_{n\in N} (A_n^2 + B_n^2) = 1$.

In this way we can choose the span Δ in advance so as to make the error in calculating M_1^* as small as required.

13. Comparison with the moving average approach

The most popular simulation technique used in recent times has been the moving average approach (see e.g. Journel (1974), Bíly and Bukoveczky (1976) and Journel and Huijbregts (1978)). A brief (and somewhat improved) presentation follows.

The method simulates a real, zero-mean Gaussian stationary stochastic process $X(t)$ with given covariance function $R(t)$ at a discrete sample of points $\{t_n : n\in(-\infty, +\infty)\}$ distant $1/2W$ from each other. Let $X(t_n) = X_n$ and $R(t_n) = r_n$. The simulation attempts to represent λ_n as a moving average of independent standard normal variables $\{V_n : n\in(-\infty, +\infty)\}$, in the form

$$X_n = \sum_{r=-\infty}^{+\infty} C_r V_{n-r} \qquad (4)$$

where the C_n are coefficients to be determined. It turns out to be convenient to set

$$C_{-n} = C_n.$$

Then $$r_n = E\{X_k X_{k+n}\} = \sum_{r=-\infty}^{+\infty} C_r C_{r+n}.$$

Let now $f_D(\omega) = \sum_{n=-\infty}^{+\infty} r_n \exp(i\omega n)$ (which is necessarily ≥ 0) be the power spectrum of X_n,

and let

$$C(\omega) = \sum_{n=-\infty}^{+\infty} C_n \exp(i\omega n).$$

Then

$$f_D(\omega) = C(\omega)C(-\omega).$$

However, from the assumption $C_{-n} = C_n$, we conclude that

$$C(\omega) = C(-\omega) = C_0 + 2\sum_{n=1}^{\infty} C_n \cos\omega n.$$

It follows that $C(\omega)$ can be taken as $\sqrt{f_D(\omega)}$, and we then have, from well-known Fourier results

$$C_n = \frac{1}{\pi}\int_0^{\pi} \sqrt{f_D(\omega)}\, \cos\,\omega n\, d\omega, \quad n = 0,1,2,\ldots \quad . \qquad (5)$$

With the C_n calculated from (5), X_n given by (4) has a covariance function exactly equal to that of $X(t)$ at all points t_n.

Note that the variance of X_n is given by $R(0) = \sum_{n=-\infty}^{\infty} C_n^2 = C_0^2 + 2\sum_{n=1}^{\infty} C_n^2$.

To obtain an approximation to X_n involving only a finite number of random variables we choose a subset of integers N such that

$$\sum_{n\in\bar{N}} C_n^2 << \sum_{n\in N} C_n^2 ,$$

where $\bar{N}$ is the complement of N in the set of non-negative integers.

We can then write

$$X_n = W_n + W_n'$$

where

$$W_n = \sum_{r\in N} C_r X_{n-r}$$

and

$$W_n' = \sum_{r\in\bar{N}} C_r X_{n-r}.$$

For fixed n, W_n and W_n' are independent. However for $n \neq m$ W_n and W_m' are correlated, unlike the decomposition of trigonometric polynomials. This approach is in a certain sense the dual of the approach advocated in this paper, in that the trigonometric polynomial represents the original process over a finite time interval $(-T, +T)$, resulting in the spectrum being discrete, while in the moving average process the time parameter is discrete, resulting in a spectrum which is band-limited.

One advantage of the moving average process is that the simulation can be continued for as long an interval of time as required, making it particularly suitable for real-time simulation.

On the other hand, the trigonometric polynomial offers the advantage of an analytical expression in closed form, which makes it eminently suitable for applications where e.g. differentiation and integration of the process

must be performed. An example of the power of this approach has been given in Section 12. No similar deterministic error estimation can be performed with the moving average process.

But the main advantage of the trigonometric polynomial approach lies in its ability to deal with narrow band processes. A perusal of Section 5 shows that for fixed T, the number of terms of the polynomial decreases when the band becomes narrower. On the other hand, when the band becomes narrower, the covariance function decays more slowly, and more terms are needed in the moving average approach.

This property of the moving average can be illustrated by calculating the $C_n'S$ for the example given in Section 5. For points distant $2W$, we find that approximately

$$C_0 = C_0\{\exp(-\tfrac{1}{2}\sigma_0^2 n^2)\}\cos n\mu_0 ,$$

where $\sigma_0 = \sqrt{2}\sigma/2W$ and $\mu_0 = \lambda_0/2W$.

It is clear that the smaller σ, the larger the number of required coefficients.

For example, it we take $2W = 200$ (a number which was actually used in Section 9) and $\sigma = 2\pi$, it turns out that we need at least 80 coefficients (as compared with 20 coefficients for the trigonometric polynomial), and, of course, for each simulation of $X(t)$ for 200 points, we need $200 + (80 \times 2) = 360$ simulations of independent standard normal variables, as against only 40 for the trigonometric polynomial. It is true that for the trigonometric polynomial we need to calculate a large number of sine and cosine values, but these can be efficiently calculated by using Fast Fourier Transform techniques.

A list of approximate values of C_n is given in Table 2 for the values given in the above paragraph.

TABLE 2

Moving Average Coefficients for X(t)

$C_0 = 0.2238$

n	c_n	n	c_m
10	-0.1434	60	2.72×10^{-8}
20	-9.856×10^{-9}	70	-0.00122
30	0.0651	80	3.53×10^{-4}
40	-0.0461	90	-8.30×10^{-5}
50	0.0135	100	-3.70×10^{-8}

14. Acknowledgements: The research embodied in this paper was begun while the author was a Visiting Professor at the Center for Stochastic Processes, University of North Carolina, Chapel Hill, and has been partially supported by the Air Force Office of Scientific Research Contract No. F49620 82 C 0009.

The author is indebted to the referee for valuable suggestions and in particular for suggesting the comparison carried out in Section 13.

REFERENCES

1. BÍLY, M. and ČAČKO, J. (1982) "Simulation of random processes with various probability density functions". J. Sound and Vibration, 81 (3) pp. 393-403.

2. BÍLY, M. and BUKOVECZKY, J. (1976) "Digital simulation of environmental processes with respect to fatigue". J. Sound and Vibration 49(4), pp. 551-568.

3. DAVIS, P.J. and RABINOWITZ, P. (1975) "Methods of numerical Integration". Academic Press, N.Y.

4. DEAK, I. (1980) "Three Digit Accurate Multiple Normal Probabilities". Numer. Math. 35, pp. 369-380.

5. HASOFER, A.M. (1982) "Simple trigonometric models for narrow-band stationary processes". Essays in Statistical Science. J. Appl. Prob. Special Volume 19A, pp. 333-344.

6. JOURNEL, A. (1974) "Simulations conditionnelles-théorie et pratigue". Doctoral Thesis C.G. Fontainebleau.

7. JOURNEL, A.B. and HUIJBREGTS, (1978) "Mining Geostatistics". Academic Press, N.Y.

8. KROPAČ, O. (1981) "A unified model for non-stationary and/or non-Gaussian random processes". J. Sound and Vibration 79(1), pp. 11-21.

9. LYON, R.H. (1970) "Statistics of combined sine waves". J. Accoustical Society of America 48, pp. 145-149.

10. MALYSHEV, I.A. and PALAGIN, Yu. I. (1981) "Mathematical modelling of multidimensional random fields". Engineering Cybernetics 6, pp. 9-17.

11. MIHAILOV, G.A. (1978) "Numerical construction of a random field with a given spectral density". Soviet Math. Dokl. 19(1), pp. 99-101.

12. MORAN, P.A.P. (1984) "The Monte Carlo evaluation of orthant probabilities for multivariate normal distributions" Aust. J. of Statistics, Vol. 26, No. 1, pp. 39-44.

13. RICE, S.O. (1954) "Mathematical analysis of random noise" (in selected papers on Noise and Stochastic Processes, Ed. N. Wax, New York: Dover Publications Inc.).

14. ROSENBLUETH, E. (1975) "Point estimates for probability moments". Proc. Nat. Acad. Sci. U.S.A. Vol. 72, No. 10, pp. 3812-3814.

15. RUBINSTEIN, R.Y. (1981) "Simulation and the Monte Carlo method". Wiley, N.Y.

15. SHINOZUKA, M. and JAN, C.-M. (1972) "Digital simulation of Random Processes and its applications". J. Sound and Vibration 25(1), pp. 111-128.

16. YANG, J.-N. (1972) "Simulation of random envelope processes". J. Sound and Vibration 21(1), pp. 73-85.

17. YANG, J.-N. (1973) "On the normality and accuracy of simulated random processes". J. Sound and Vibration 26(3), pp. 417-428.

END

1-87

DTIC